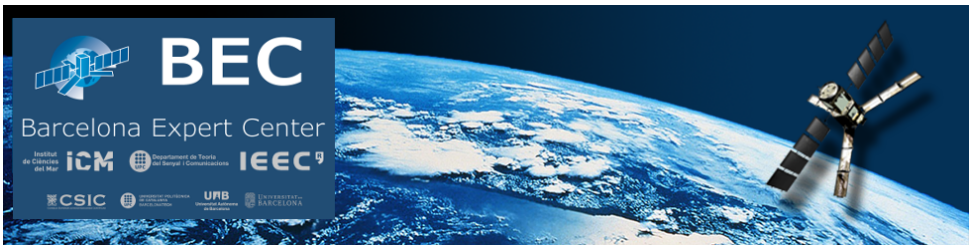


File: BEC-SMOS-0003-PD-SSS-MED.pdf , version 1.0

Title: SMOS-BEC Mediterranean Region SSS Product Description.

Authors: BEC Team.

Contact: smos-bec@icm.csic.es

Date: 02/08/2019

SMOS-BEC MEDITERRANEAN REGION SSS PRODUCT DESCRIPTION

Abstract: This technical note describes the Debiased non-Bayesian retrieval of SMOS SSS for the Mediterranean and North Atlantic region distributed by the BEC team through its data visualization and distribution service http://bec.cmima.csic.es

Contents

1 INTRODUCTION

A new methodology using a combination of debiased non-Bayesian retrieval, DINEOF (Data Interpolating Empirical Orthogonal Functions) and multifractal fusion has been used to obtain Soil Moisture and Ocean Salinity (SMOS) Sea Surface Salinity (SSS) fields over the North Atlantic Ocean and the Mediterranean Sea. The debiased non-Bayesian retrieval mitigates the systematic errors produced by the contamination of the land over the sea. Besides, this retrieval improves the coverage by means of multiyear statistical filtering criteria. This methodology allows obtaining SMOS SSS fields in the Mediterranean Sea. However, the resulting SSS suffers from a seasonal (and other time-dependent) bias. This time-dependent bias has been characterized by means of specific Empirical Orthogonal Functions (EOFs). Finally, high resolution Sea Surface Temperature (OSTIA SST) maps have been used for improving the spatial and temporal resolution of the SMOS SSS maps. The presented methodology practically reduces the error of the SMOS SSS in the Mediterranean Sea by half. As a result, the SSS dynamics described by the new SMOS maps in the Algerian Basin and the Balearic Front agrees with the one described by in situ SSS, and the mesoscale structures described by SMOS in the Alboran Sea and in the Gulf of Lion coincide with the ones described by the high resolution remotely-sensed SST images (AVHRR).

2 SCIENCE ALGORITHM OVERVIEW

2.1 Level 2 Algorithms

The debiased non-Bayesian retrieval of SSS introduced in [Olmedo et al., 2017] aims to correct two known issues: the systematic biases caused by the presence of land masses and radio interference, and the data gaps due to the non-convergence of the retrieval algorithm.

2.1.1 Input data

The Brightness temperatures (TBs) obtained from SMOS MIRAS L1B TBs v620 provided by ESA are used as an input for the SMOS SSS retrieval. The L1B v620 product contains the Fourier coefficients of the measured brightness temperature. Starting from this product, using ESA's Earth Observation Customer Furnished Item (EOCFI) orbit propagation libraries [**?**] and following a similar procedure as the one used in the operational SMOS level 1 processor chain [**?**], the measured TBs are obtained in the antenna reference frame (ARF). The unique difference from the standard processor is the number of points contained per snapshot (*i.e.* the resolution). The operational processor uses, at antenna level, a hexagonal grid of 128×128 points (*i.e.* $2^7 \times 2^7$). The projection of this antenna grid into the ground provides a resolution of about 15 km at bore-sight. This resolution is more than twice SMOS theoretical finer resolution [McMullan et al., 2008]. We have thus reduced the computational cost without actually losing information by using an antenna hexagonal grid of 64×64 ($2^6 \times 2^6$) points.

2.1.2 Forward Model

The forward model linking the modeled TB to SSS relies on the dielectric constant model proposed by Klein and Swift [**?**] which non-linearly depends on SSS and SST. Nevertheless, the measured TB not only contains information about brightness temperature of the flat sea, but also contributions due to other main sources: the roughness of the sea surface [**?**], the reflected emission of the atmosphere, the reflection on the sea surface of the galactic emission ([**?**]) and the sun glitter [**?**]. Therefore, all these additional contributions to the TB(hj, SSS, p1, . . . , pNp) term must be modeled at the bottom of the atmosphere (BOA) and then translated to the ARF prior to minimizing Eq. (1). Thus, the atmospheric attenuation effect over brightness temperature must be taken into account together with the direct emission of the atmosphere itself to estimate the modeled TB at the top of the atmosphere (TOA). Finally, to go from TOA to ARF, the ionosphere must be taken into account. The ionospheric effect translates into a rotation in the polarization components of TB. All these contributions are described in detail in Ocean SMOS Team (2016).

2.1.3 Non-Bayesian retrieval of SSS

We retrieve a single SSS value for each TB measurement, that is, along the same dwell line we have a value of SSS for each valid incidence angle, namely:

$$F^{j}_{non-Bayesian}(SSS) = \left[I^{meas}(\theta_j) - I(\theta_j, SSS, p_1, \ldots, p_{N_p})\right]^2, \quad (1)$$

where the super index j indicates one of the N_m available incidence angles. The term $I = (TB_v + TB_h)/2$, both for the forward modeled (section 2.1.2) and measured data, is the First Stokes parameter at BOA divided by 2; by summing up vertical (TB_v) and horizontal (TB_h) polarization we obtain a term which is independent from Faraday rotation and simplifies its processing, although the retrieval could be done with each polarization independently. For the optimization, the other geophysical

variables are given a fixed value, that of the geophysical priors p_k^0.

2.1.4 SSS spatial bias characterization: definition of a SMOS-based climatology

The characterization of bias is based on a classification of the non-Bayesian single-angle salinities retrieved using eq. (1). A first geophysical consistency filter on non-Bayesian salinities is applied: we discard any value out of the range $[0, 50]$. Then, the single-angle SSS values are grouped together according to their geolocation (latitude and longitude in a cylindrical grid of $0.25^o \times 0.25^o$, φ and λ), overpass direction (ascending or descending, denoted by a binary variable d), across-track distance to the center of swath (in 50-km bins, denoted by x) and incidence angle (in 5° bins, denoted by θ). For each given 5-tuple, $c = (\varphi, \lambda, d, x, \theta)$, we take all the retrievals $\{SSS(\varphi, \lambda, d, x, \theta)\}$ in the period we use (2011 to 2016) and construct the associated histograms (see Figure 1). The choice of the parameters used in the definiton of the tuple is further detailed in [Olmedo et al., 2017] (section 2.2.2).

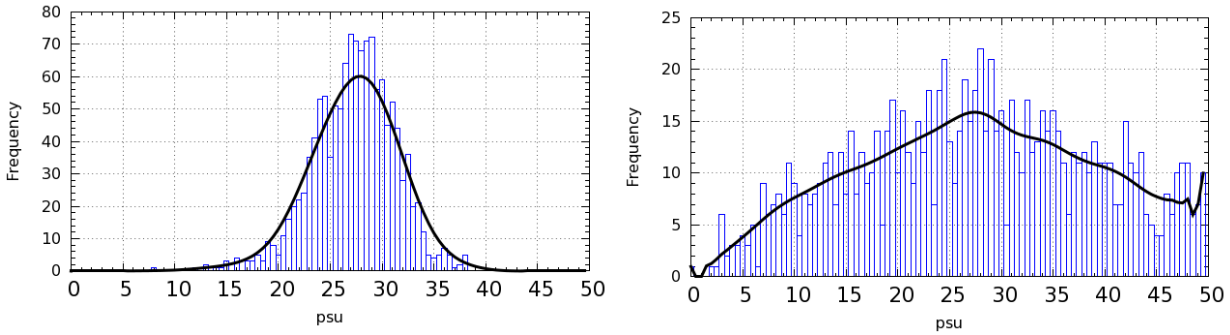


Figure 1: SMOS SSS histogram associated with $c = (20, -29, A, 225, 52)$ (left) and to $c = (31, 133, A, 275, 47)$ (right) (see the text for more details).

For every tuple, the corresponding SSS histogram is defined with bins of 0.5 psu. In order to improve the accuracy of the histogram and minimize the dependency on the histogram discretization, the mean of SSS inside each bin is considered as the representative value of SSS on that bin. The mode of the histograms are taken as a measure of the central reference value; the advantage of taking the mode, instead of other statistics, is that the mode is unaffected by the presence of outliers or by the skewness of the distribution.

The accuracy of the estimation of the mode of any histogram is relatively low, mainly because of the lack of sampling. To overcome this issue we have applied three times a weighted averaging window (with a size of seven points, including the central point, the three points to its left and the three points to its right) to each histogram to eliminate statistically non-significant fluctuations. As shown in Figure 1 (black line), this leads to a better determination of the location of the maximum probability (i.e., the mode). The resulting smoothed histogram is only used for estimating the mode; the rest of statistical parameters are computed from the original (not-smoothed) histogram. We compute a SMOS-based climatological value for each given 5-tuple (denoted as $sss\_clim(\varphi, \lambda, d, x, \theta)$) by averaging all the retrieved single-angle SSSs lying in a range of $\pm\sigma$ around the estimated mode, where σ is the standard deviation of the single-angle SSSs for that 5-tuple.

Figure 2 shows two maps of the SMOS-based climatologies, corresponding to ascending overpasses, $x = 0$ km, $\theta = 5°$ (top) and $\theta = 35°$ (bottom). Significant differences are found for different θ values (that is, depending on the relative position of the pixel in the ARF), see Figure 3. The differences are notably large close to the coast due to LSC, although they can be significant also elsewhere. A close inspection of the SMOS-based climatologies shows that for any given value of x and θ the deviation

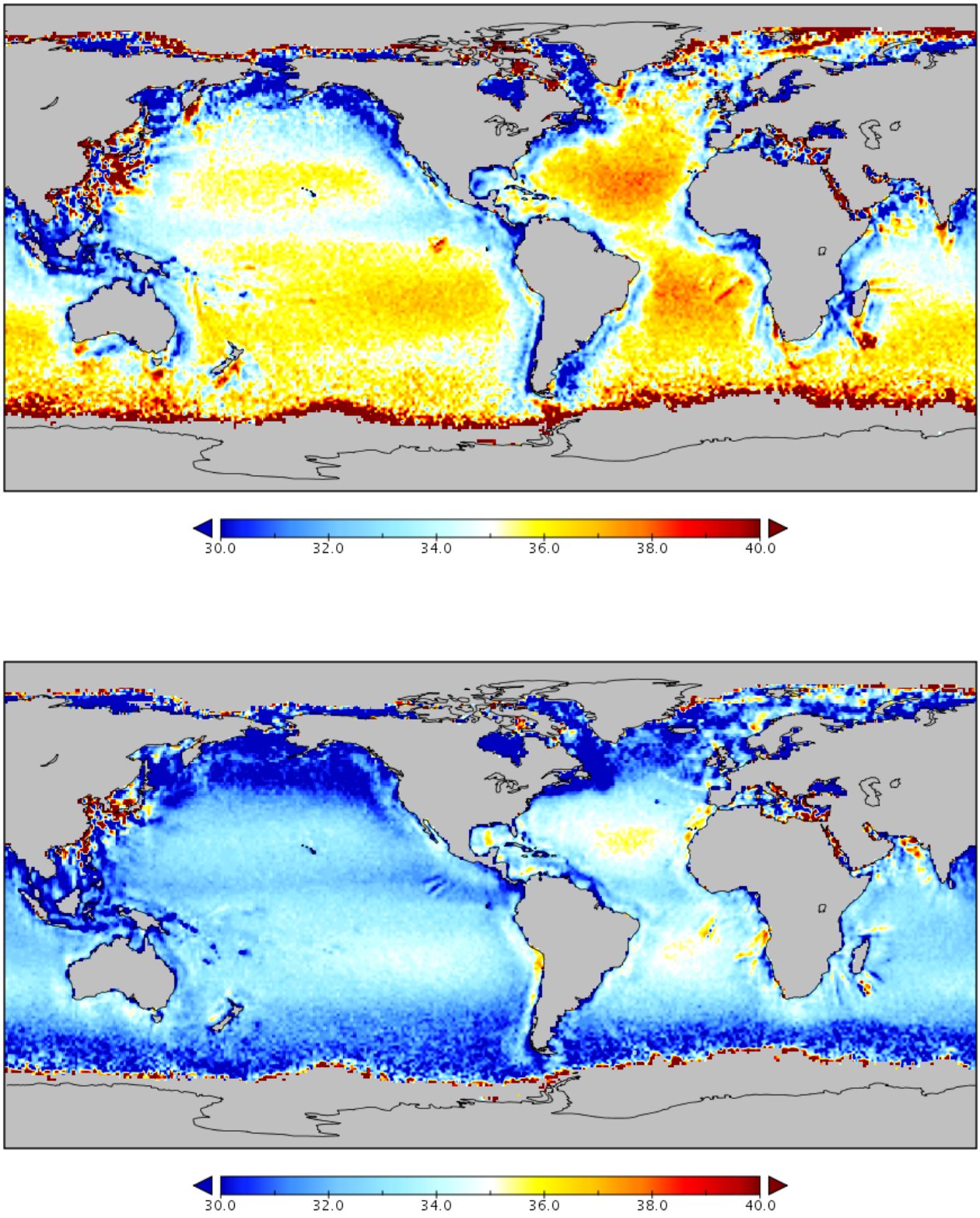


Figure 2: SMOS-based climatology for ascending overpasses and $x = 0$ km and $\theta = 5°$ (top) and $x = 0$ km and $\theta = 35°$ (bottom).

of SMOS-based climatology from a standard annual climatology is far from being spatially constant, even in open sea (see Figure 4).

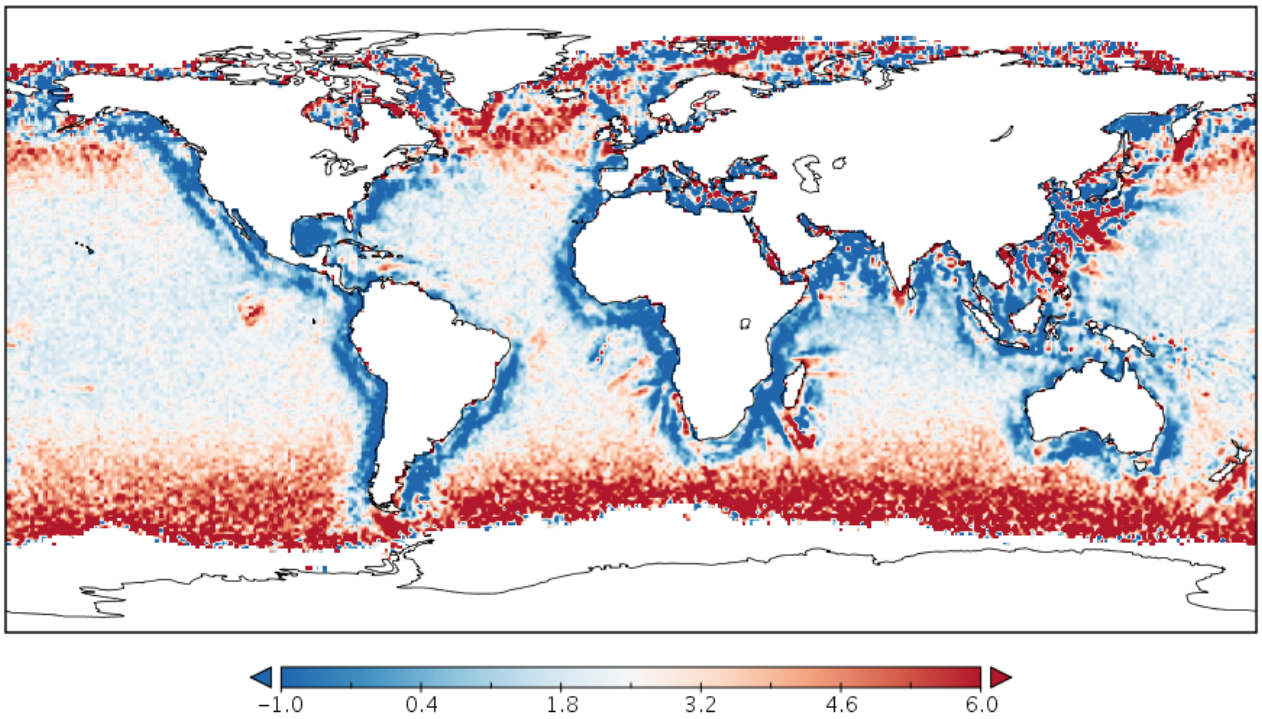


Figure 3: Map of differences between the SMOS-based climatology for ascending overpasses and x=0 km and $\theta = 5^o$ and the SMOS-based climatology for ascending overpasses and x=0 km and $\theta = 35^o$.

2.1.5 Spatial bias correction

For the generation of higher-level SSS products, single-angle SSS retrievals (Section 2) are corrected with their corresponding SMOS-based climatological value (sss_{clim}), thus creating a set of so-called SMOS-based anomalies (sss_{an}):

$$sss_{an}(\varphi, \lambda, d, x, \theta) = SSS(\varphi, \lambda, d, x, \theta) - sss_{clim}(\varphi, \lambda, d, x, \theta). \quad (2)$$

In order to obtain an absolute value of SSS, a time-independent reference SSS value must be added to the L2 product. We use here the annual SSS climatology provided by the World Ocean Atlas 2013 (WOA2013) at $0.25^o \times 0.25^o$ (average decadal product, which is accessible at [**?**] [Zweng et al., 2013]).

2.1.6 SSS filtering criteria for non-Bayesian approach

After computing SMOS-based climatologies, several quality control criteria are applied to discard poor quality values. For a given 5-tuple, $c = (\varphi, \lambda, d, x, \theta)$, a SMOS-climatological value is discarded (and hence all related single-angle SSS retrievals) when the associated histogram suffers from one or more of the following conditions:

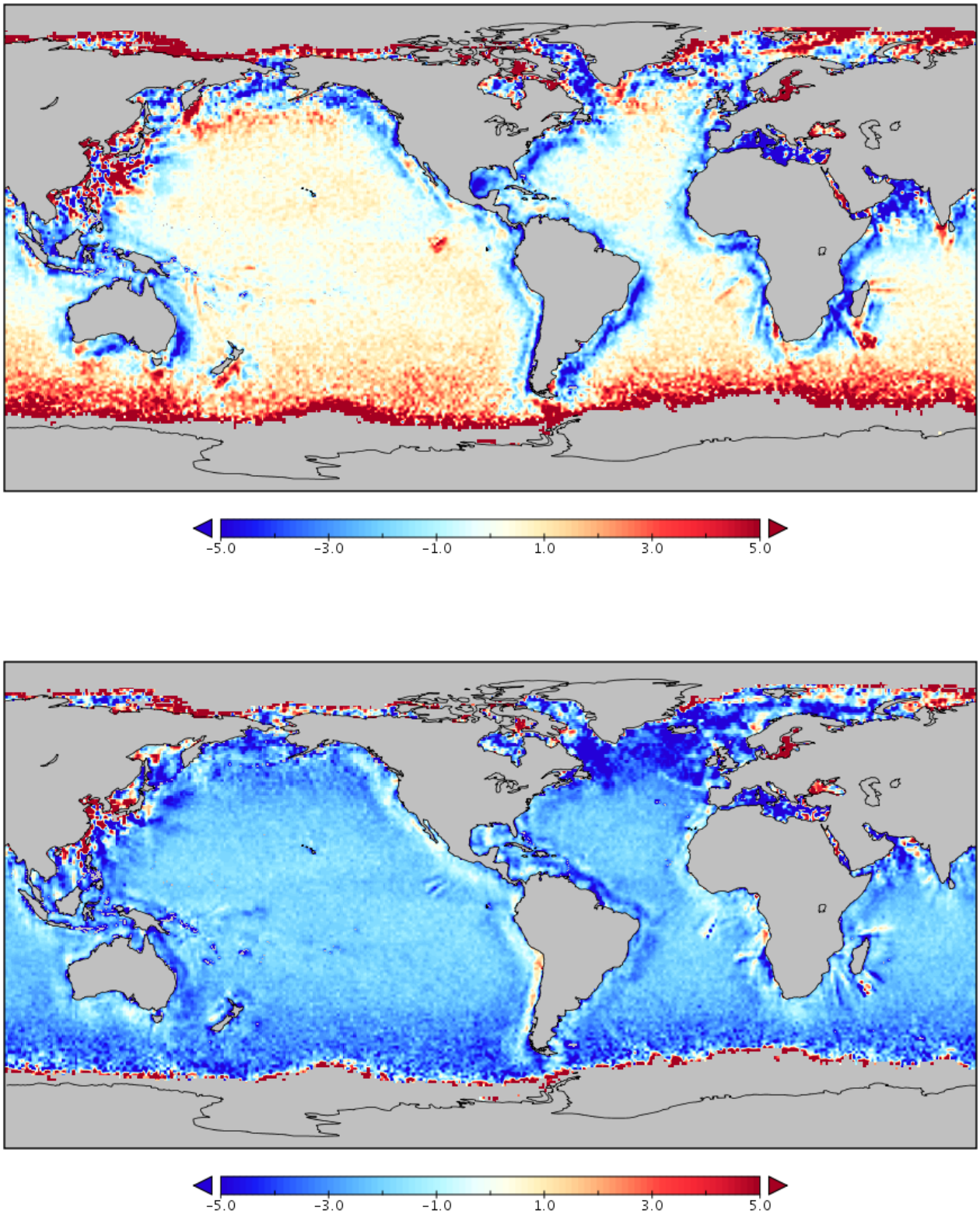


Figure 4: Difference between SMOS-based climatology and the annual WOA2013 climatology for ascending overpasses and $x = 0$ km and $\theta = 5°$ (top) and $x = 0$ km and $\theta = 35°$ (bottom).

- It has less than 100 measurements;
- Its standard deviation is greater than 10 psu;
- The absolute value of its normalized skewness is greater than 1;
- Its kurtosis is lower than 2

Two additional filtering criteria are applied to the SMOS-based anomalies:

- If the histogram corresponding to a given geographical and orbital coordinates $c = (\varphi, \lambda, d, x, \theta)$ has been discarded following the criteria of Section 2.1.6, then all the $sss_{an}(\varphi, \lambda, d, x, \theta)$ are also discarded.
- When the SMOS-based anomaly is greater than the standard deviation of the distribution, the value is considered an outlier and, therefore, it is also discarded.

As an example, see the two histograms displayed in Figure 1. The one in the left is accepted, as none of the conditions above is verified (it has 581 measurements, a standard deviation of 2.31, skewness equal to 0.97 and kurtosis equal to 3.74). The one in the right has 369 measurements, a standard deviation 5.82 and a 0.19 skewness, which are right, but its kurtosis (1.79) is too low and, therefore, it is discarded.

2.2 Level 3 Algorithms

2.2.1 Mitigation of time-dependent biases: removal of EOFs from the DINEOF basis

In [Olmedo et al., 2017] the mitigation of the time-dependent biases is done by removing of the mean value of the SMOS-based anomaly of each 9-day map. This time-dependent bias correction provides quite accurate global maps (as shown in [Olmedo et al., 2017]). However, in some specific more challenging regions, as the Mediterranean Sea, some residual time-dependent biases still remain (as observed in [Isern-Fontanet et al., 2016]). Therefore, for this regional product we do not apply this time correction to our 9-day maps. Instead, we try to remove the residual biases by means of Data Interpolating Empirical Orthogonal Functions (DINEOF).

DINEOF (Data Interpolating Empirical Orthogonal Functions, [Beckers and Rixen, 2003, Alvera-Azcárate et al., 2
has been used to analyze the L3 SSS binned fields obtained using the debiased non-Bayesian approach. DINEOF is a technique to reconstruct missing data and reduce noise in geophysical data sets using an EOF basis. The most dominant EOFs, calculated iteratively to overcome the presence of missing data, are used to infer the missing information. As the amount of missing data in the 9-day L3 binned SSS data set used in this work is very low (5.3%), the main aim of the use of DINEOF with these data is to reduce noise and remove non-physical patterns. A filter has been applied to the covariance matrix, following [Alvera-Azcárate et al., 2009]. This filter allows to ensure a smooth transition between subsequent images, by providing information on the time step between them which is otherwise not considered in the EOF decomposition. The filter length is of 1.98 days.

Six years (2011-2016) of 9-day binned L3 SMOS SSS maps at $0.25^o \times 0.25^o$ have been decomposed into 30 EOFs. Our spatial domain contains part of the North Atlantic Ocean and the full Mediterranean

Sea. Let us call our spatial domain $\Omega = [26^o, 50^o]N \times [50^oW, 38^oE]$. Thus, the initial binned SMOS SSS ($SSS_{bin}(t, \varphi, \lambda)$) has been approximated by:

$$SSS_{bin}(t, \varphi, \lambda) \approx SSS_{mean}(t, \varphi, \lambda) + \sum_{i=1}^{30} s(i) * v(t, i) * u(\varphi, \lambda, i), \tag{3}$$

where φ, λ represent the latitude and longitude respectively, t represents the time variable, $v(t, i)$ is the time-dependent part of the EOF decomposition, $u(\varphi, \lambda, i)$ is the spatial-dependent part of the EOF decomposition and $s(i)$ are the singularity values. For the same time period (2011-2016), we have collocated Argo float SSS data with the SMOS SSS maps. Then, at every day, t_0, we have a set of locations with Argo and SMOS SSS. Let us call $\Theta(t_0) \subset \Omega$, the set of locations where Argo and SMOS SSS have acquisitions at time t_0. We define the following time series:

$$\epsilon_0(t) = \sum_{\varphi,\lambda \in \Theta(t)} (SSS_{bin}(t, \varphi, \lambda) - SSS_{Argo}(t, \varphi, \lambda)) \tag{4}$$

If we assume that Argo covers the domain uniformly, this time series shows the time-dependent bias of the SMOS SSS. In particular, we can approximate it by:

$$\epsilon_0(t) \approx \sum_{\varphi,\lambda \in \Theta(t)} \left\{ \left(SSS_{mean}(t, \varphi, \lambda) + \sum_{i=1}^{30} s(i) * v(t, i) * u(\varphi, \lambda, i) \right) - SSS_{Argo}(t, \varphi, \lambda) \right\} \tag{5}$$

We want to remove the EOFs which describe the most part of the time-dependent bias. So we compute the correlation between $\epsilon(t)$ and each one of the 30 time series corresponding to:

$$\omega_i(t) = \sum_{\varphi,\lambda \in \Theta(t)} s(i) * v(t, i) * u(\varphi, \lambda, i) \tag{6}$$

Such that each ω_i corresponds to the contribution of the $i-th$ EOF to the SMOS SSS at the collocated locations with Argo SSS. Therefore, the $\omega_i(t)$ with the highest correlation with $\epsilon_0(t)$ will describe the highest part of the residual time-dependent bias of the initial SMOS SSS data. Let us call i_0, the EOF which provides the highest correlation with $\epsilon_0(t)$. We remove the EOF i_0 from the initial approximation and redefine:

$$\epsilon_1(t) = \epsilon_0(t) - \omega_{i_0}(t) \tag{7}$$

Then we iterate the process. We compute the correlation of all the $\omega_i(t)$ (except for $\omega_{i_0}(t)$, which has been removed) with $\epsilon_1(t)$. The $w_i(t)$, with the highest correlation will describe the highest part of the residual time-dependent bias. We iterate until, the correlation is not significant, or when the residual $\epsilon_n(t)$ is too small. In our case, as we are interested in the characterization and correction of the seasonal and other eventual time-dependent biases, we expect that the signal of this error will be concentrated in a few EOFs. The final selection of the number of EOFs to be removed has been done by considering the partial solutions obtained in each iteration (and shown in Table 1 and Figure 5). We explain it in more detail below.

In Table 1 the different iteration steps are shown. In the first step (first row), the mean value of the initial residual $\epsilon_0(t)$ is -0.07 PSU (fourth column) and the corresponding standard deviation (fifth column) is 0.23 PSU. The EOF which better correlates with ϵ_0 is the EOF number 2 (second column) with a correlation value of 0.69 (third column). The first plot of Figure 5 shows the time evolution of $\epsilon_0(t)$ (green squares) and $\omega_2(t)$ (violet triangles). In the second row of Table 1 the results of the second iteration are shown: the mean value of the residual $\epsilon_1(t) = \epsilon_0(t) - \omega_2(t)$ and the corresponding standard deviation are -0.04 PSU and 0.17 PSU (respectively) (fourth and fifth column of the second

Table 1: Results of the first 10 iterations from the time-dependent bias deletion from EOF decomposition. The first column shows the residual remaining in each iteration. In the second column the number of the EOF that provides the largest correlation with the residual in each step. The correlation between the different EOFs contributions to the SMOS SSS and the time-dependent bias is shown in the fifth column. In the third and fourth column the mean and standard deviation (respectevely) of the residual error before the correction.

Residual	EOF with largest correlation	Correlation	Mean	Std
ϵ_0	2	0.69	-0.07	0.23
$\epsilon_1 = \epsilon_0 - \omega_2$	1	0.46	-0.04	0.17
$\epsilon_2 = \epsilon_1 - \omega_1$	6	0.27	-0.03	0.15
$\epsilon_3 = \epsilon_2 - \omega_6$	3	0.28	-0.05	0.15
$\epsilon_4 = \epsilon_3 - \omega_3$	5	0.31	-0.04	0.14
$\epsilon_5 = \epsilon_4 - \omega_5$	7	0.21	-0.05	0.14
$\epsilon_6 = \epsilon_5 - \omega_7$	8	0.20	-0.05	0.14
$\epsilon_7 = \epsilon_6 - \omega_8$	4	0.19	-0.05	0.13
$\epsilon_8 = \epsilon_7 - \omega_4$	27	0.17	-0.05	0.13
$\epsilon_9 = \epsilon_8 - \omega_{27}$	9	0.16	-0.05	0.13

row of Table 1). The EOF with the largest correlation with $\epsilon_1(t)$ is the EOF number 1. In this case, the correlation between $\omega_1(t)$ and $\epsilon_1(t)$ is 0.46. The second plot of Figure 5 shows the time evolution of $\epsilon_1(t)$ (green squares) and $\omega_1(t)$ (violet triangles)). In the third iteration, the mean value and standard deviation of the residual $\epsilon_2(t) = \epsilon_1(t) - \omega_1(t)$ are -0.03 PSU and 0.15 PSU (respectively). The EOF with the largest correlation (0.27) is the EOF number 6. The third plot of Figure 5 shows the time evolution of $\epsilon_2(t)$ and the $\omega_6(t)$.

We have removed only these three EOFs. As observed in Table 1, from the third iteration onward (the table only includes the first 10 steps), the resulting residuals are almost the same, and they do not improve the correlation.

2.2.2 Reduction of white noise by means of objective analysis

We have applied a scheme of objective analysis to the resulting debiased and time-corrected L3 maps. The scheme is the one described in [Olmedo et al., 2017] which is also the same applied for the generation of World Ocean Atlas ([Zweng et al., 2013]). The algorithm consists of averaging the binned SSS fields in three iterative steps, such that in each step a different radius of influence is considered. In our case the three different radii are 175, 125 and 75 km respectively.

2.3 L4 Algorithm

2.3.1 Improving the spatial and temporal resolution by mean of multifractal fusion

It has been shown that different ocean scalars possess a multifractal structure which is due only to the action of the underlying ocean currents [Turiel et al., 2008] and thus is shared by different scalars [Isern-Fontanet et al., 2007, Nieves et al., 2007]. It has been shown that this degree of redundancy among different ocean scalars can be used to enhance the quality of one noisy ocean variable if a different, less noisy variable is known: this is the so-called multifractal fusion ([Umbert et al., 2014]). Multifractal fusion has been shown to be useful to increase the spatial- and temporal- resolution

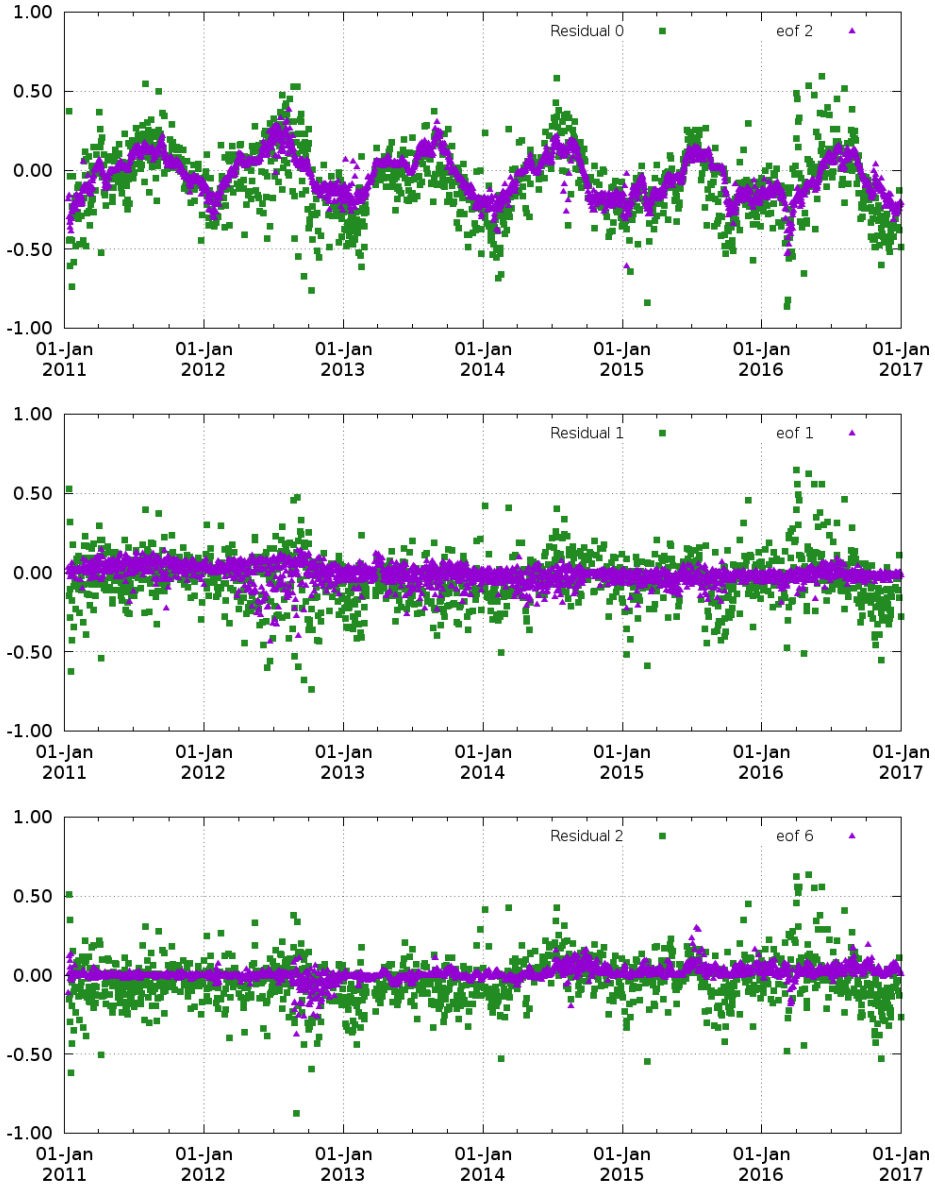


Figure 5: Time evolution of the mean difference between SMOS and Argo SSS ($\epsilon(t)$ green squares) and the contribution to the SSS of the EOF which provides the largest correlation with the residual ($\omega(t)$ violet triangles). On the top the results corresponding to the first iteration step; in the middle the ones corresponding to the second iteration; and in the bottom the results of the third and last iteration

of ocean scalars (particularly, SMOS SSS) while, at the same time, reducing the level of noise ([Olmedo et al., 2016] in case of SMOS, and [Hernández-Carrasco et al., 2015] in case of CO_2 satellite imagery).

The main idea of the algorithm is to assume that we have two different ocean variables measured by remote sensing means such that one of them, θ is considerably less noisy than the other, s. The main goal of the fusion method is to use the more reliable information provided by θ to reduce the noise and improve the spatial structure of s. In our case, this is done by assuming that both, θ and s, have the same singularity exponents (SE).

Now, if both variables have the same SE, a relation must exist between the spatial variations of both variables because they behave the same way under changes of scale (they are ruled by the same SE). As discussed in [Umbert et al., 2014], the correspondence between SE must be smooth as it cannot be an additional source of singular behavior. In particular, the relation between the gradient of s and that of θ can be expressed in terms of a smooth matrix Φ:

$$\nabla s(\vec{x}) = \Phi(\vec{x})\nabla\theta(\vec{x}) \tag{8}$$

The estimation of the matrix $\Phi(\vec{x})$ from equation (8) is not straightforward: this problem is ill-posed, as for each x the number of unknowns (the elements of the matrix $\Phi(\vec{x})$) is d^2 while the number of equations is d.

In [Olmedo et al., 2016], different strategies to estimate this matrix from real data are investigated, with the only constraint that the matrix $\Phi(\vec{x})$ must smoothly depend on $\vec{x}$. We use here the scalar approach which consists in assuming that is proportional to the identity matrix, i.e., the matrix can be simplified to a scalar a, as in [Umbert et al., 2014]:

$$\nabla s(\vec{x}) = a(\vec{x})\theta(\vec{x}) \tag{9}$$

This expression can be integrated under the assumption that the gradient of a is negligible as compared to the gradients of s and θ namely

$$s(\vec{x}) = a(\vec{x})\theta(\vec{x}) + b(\vec{x}) + \varepsilon(\vec{x}) \tag{10}$$

where the functions a, b have small gradients and $\varepsilon(\vec{x})$ is the additive error to be minimized.

3 MEDITERRANEAN SSS PRODUCT

3.1 Ocean files structure

The resulting Level 3 and Level 4 products are distributed in netCDF format and the name of each file follows the layout:

```
BEC_AAAAAA_B_CCCCCCCCCCCCCCC_DDDDDDDDDDDDDDD_EEEEEEE_FFF_GGG.nc
```

Where each field of the filename is as follows:

- `AAAAAA`: is the product's name:
  - `BINNED`: Binned product
  - `OI____`: Optimal Interpolation product
  - `L4_SST`: Fused product using singularity analysis techniques derived from SST
  - `EXPSST`: Singularity Exponents
- `B`: Indicates the orbit composition of the product.
  - `A` for products composed by ascending orbits
  - `D` for products composed by descending orbits
  - `B` for products composed by both types of orbits
- `CCCCCCCCCCCCCCC`: Starting UTC time (YYYYMMDDThhmmss) of the first L2 product used to create the L3/L4 product. This is an inherited value in products not derived directly from Level 2 orbits.
- `DDDDDDDDDDDDDDD`: Ending UTC time (YYYYMMDDThhmmss) of the last L2 product used to create the L3/L4 product. This is an inherited value in products not derived directly from Level 2 orbits (Optimal Interpolation and L4 products).
- `EEEEEEE`: Internal code that designates the filtering applied. This is an inherited value in products not derived directly from Level 2 orbits.
- `FFF`: Grid size of the product in a lat-lon grid multiplied by 100
- `GGG`: Version number of the file starting at 001

3.2 Data Definition

The structure of the netCDF files with its dimensions, variables and global attributes is detailed here:

3.2.1 L3

```
dimensions:
    lat = 96 ;
    lon = 352 ;
```

```
     time = UNLIMITED ; // (1 currently)
variables:
     float lat(lat) ;
        lat:standard_name = "latitude" ;
        lat:long_name = "latitude" ;
        lat:units = "degrees_north" ;
        lat:axis = "y" ;
        lat:missing_value = -999.f ;
        lat:FillValue = -999.f ;
     float lon(lon) ;
        lon:standard_name = "longitude" ;
        lon:units = "degrees_east" ;
        lon:long_name = "longitude" ;
        lon:axis = "x" ;
        lon:missing_value = -999.f ;
        lon:FillValue = -999.f ;
     float oa_sss(time, lat, lon) ;
        oa_sss:standard_name = "objective_analyzed_sea_surface_salinity" ;
        oa_sss:long_name = "Objective Analyzed Sea Surface Salinity" ;
        oa_sss:units = "psu" ;
        oa_sss:description = "sea surface salinity objective analyzed using debiased non-Bayes
and DINEOF" ;
        oa_sss:missing_value = -999.f ;
        oa_sss:_FillValue = -999.f ;
     int time(time) ;
        time:standard_name = "time" ;
        time:long_name = "time" ;
        time:units = "seconds since 1970-1-1 00:00:00" ;
        time:time = "t" ;
        time:coordinate_defines = "center" ;
        time:calendar = "gregorian" ;
        time:_CoordianteAxisType = "Time" ;
global attributes:
     :title = "Sea Surface Salinity L3 map" ;
     :date_created = "2017-09-22 08:14:01 GMT" ;
     :geospatial_lon_units = "degrees_east" ;
     :geospatial_lat_units = "degrees_north" ;
     :institution = "SMOS Barcelona Expert Centre, ICM-CSIC / UPC, Barcelona, Spain; GHER
- University of Liege, Belgium" ;
     :copyright = "BEC/GHER research products are freely distributed.  SSS have been retrieved
following the algorithm described in Olmedo, E. et al., De-biased non-Bayesian Retrieval:
a novel approach to SMOS Sea Surface Salinity, Remote Sensing of Environment 193 (2017)
103126.  The SSS have been time-corrected by means of multivariate analysis as described
in A. Alvera-Azcarate, et al., Analysis of SMOS sea surface salinity data using DINEOF,
Remote Sensing of Environment 180 (2016) 137-145.  Finally, fusion with SST from OSTIA
has been using for improving the spatial and temporal resolutions as described in Olmedo,
E. et al., Improving time and space resolution of SMOS salinity maps using multifractal
fusion, Remote Sensing of Environment 180 (2016) 246263.  The complete description of
the methodology as well as the analysis of the quality assessment of the product can be
```

found in Olmedo, E. et al., Improving SMOS Sea Surface Salinity in the Western Mediterranean Sea through Multivariate and Multifractal Analysis, Remote Sensing, 2018, 10(3). If these data are used for publication, the following ackowlegment should be included: These data were produced by the Barcelona Expert Centre (http://bec.icm.csic.es/) and the GHER - University of Liege, Belgium (http://labos.ulg.ac.be/gher/). The Barcelona Expert Center is a joint initiative of the Spanish Research Council (CSIC) and Technical (University of Catalonia (UPC), mainly founded by the Spanish National Program on Space." ;
:references = "Olmedo, E. et al., Improving SMOS Sea Surface Salinity in the Western Mediterranean Sea through Multivariate and Multifractal Analysis, Remote Sensing, 2018, 10(3), 485; doi:10.3390/rs10030485." ;
:Conventions = "CF-1.4" ;

3.2.2 L4

```
 dimensions:
     lat = 480 ;
     lon = 1760 ;
     time = UNLIMITED ; // (1 currently)
variables:
     float lat(lat) ;
        lat:standard_name = "latitude" ;
        lat:long_name = "latitude" ;
        lat:units = "degrees_north" ;
        lat:axis = "y" ;
        lat:missing_value = -999.f ;
        lat:FillValue = -999.f ;
        lat:comment = "Latitude geographical coordinates,WGS84 projection" ;
     float lon(lon) ;
        lon:standard_name = "longitude" ;
        lon:units = "degrees_east" ;
        lon:long_name = "longitude" ;
        lon:axis = "x" ;
        lon:missing_value = -999.f ;
        lon:FillValue = -999.f ;
        lon:comment = "Longitude geographical coordinates,WGS84 projection" ;
     float l4_sss(time, lat, lon) ;
        l4_sss:standard_name = "l4_sea_surface_salinity" ;
        l4_sss:long_name = "L4 Sea Surface Salinity" ;
        l4_sss:units = "psu" ;
        l4_sss:description = "Sea Surface Salinity analysis using DINEOF and Data Fusion
of SST" ;
        l4_sss:missing_value = -999.f ;
        l4_sss:_FillValue = -999.f ;
     int quality_flag(time, lat, lon) ;
        quality_flag:standard_name = "quality_flag" ;
        quality_flag:long_name = "Quality control flag" ;
        quality_flag:units = "-" ;
        quality_flag:description = "When the flag is equal to 1 the corresponding value
```

```
of salinity has been extrapolated." ;
        quality_flag:missing_value = 0 ;
    int time(time) ;
        time:standard_name = "time" ;
        time:long_name = "time" ;
        time:units = "seconds since 1970-1-1 00:00:00" ;
        time:time = "t" ;
        time:coordinate_defines = "center" ;
        time:calendar = "gregorian" ;
        time:_CoordianteAxisType = "Time" ;
global attributes:
    :title = "Sea Surface Salinity L3 map" ;
    :institution = "SMOS Barcelona Expert Centre, ICM-CSIC / UPC, Barcelona, Spain; GHER
- University of Liege, Belgium" ;
    :copyright = "BEC/GHER research products are freely distributed.  SSS have been retrieved
following the algorithm described in Olmedo, E. et al., De-biased non-Bayesian Retrieval:
a novel approach to SMOS Sea Surface Salinity, Remote Sensing of Environment 193 (2017)
103126.  The SSS have been time-corrected by means of multivariate analysis as described
in A. Alvera-Azcarate, et al., Analysis of SMOS sea surface salinity data using DINEOF,
Remote Sensing of Environment 180 (2016) 137-145.  Finally, fusion with SST from OSTIA
has been using for improving the spatial and temporal resolutions as described in Olmedo,
E. et al., Improving time and space resolution of SMOS salinity maps using multifractal
fusion, Remote Sensing of Environment 180 (2016) 246263.  The complete description of
the methodology as well as the analysis of the quality assessment of the product can be
found in Olmedo, E. et al., Improving SMOS Sea Surface Salinity in the Western Mediterranean
Sea through Multivariate and Multifractal Analysis, Remote Sensing, 2018, 10(3).  If these
data are used for publication, the following ackowlegment should be included:  These data
were produced by the Barcelona Expert Centre (http://bec.icm.csic.es/) and the GHER -
University of Liege, Belgium (http://labos.ulg.ac.be/gher/).  The Barcelona Expert Center
is a joint initiative of the Spanish Research Council (CSIC) and Technical (University
of Catalonia (UPC), mainly founded by the Spanish National Program on Space." ;
    :references = "Olmedo, E. et al., Improving SMOS Sea Surface Salinity in the Western
Mediterranean Sea through Multivariate and Multifractal Analysis, Remote Sensing, 2018,
10(3), 485; doi:10.3390/rs10030485." ;
```

3.3 Data Access

The Mediterranean and Atlantic SMOS SSS v1.0 produced at BEC is freely available through a SFTP server. If your browser is SFTP compatible you can browse directly from sftp://becftp.icm.csic.es:27500 address. The data can be download after completing registration in our BEC ftp service (http://bec.icm.csic.es/bec-ftp-service-registration/). For more information about the BEC ftp service, please visit http://bec.icm.csic.es/bec-ftp-service/. For any further assistance, contact to smos-bec@icm.csic.es.

The following diagram tree shows the complete path to the Mediterranean SMOS SSS v1.0 repository.

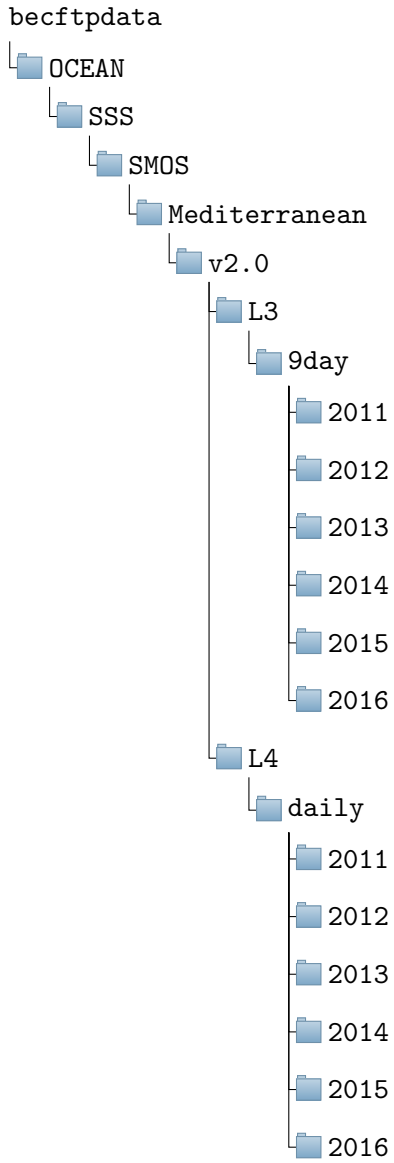


Table 2: Statistics of the differences between SMOS and Argo SSS: the mean, the standard deviation and the root mean squared. The statistics are computed over two different domains: First row: DOM which contains the Mediterranean Sea and part of the North Atlantic; and second row: MED which contains the Mediterranean Sea only.

	L3			L4		
	mean	std	rms	mean	std	rms
DOM	-0.06	0.24	0.26	-0.07	0.23	0.26
MED	-0.16	0.34	0.43	-0.19	0.29	0.39

A QUALITY ASSESSMENT

The aim of this section is to report the differences between the regional non-Bayesian debiased L3 and L4 SMOS SSS for Mediterranean distributed by BEC and other SSS in situ observations, such as: the Argo data (section A.1), mooring data (section A.2) and the TRANSMED transects (section A.3).

NOTE that the regional non-Bayesian debiased L3 and L4 (fused) SMOS SSS for Mediterranean distributed by BEC include a time-dependent bias correction by means of Data Interpolating Empirical Orthogonal Functions (DINEOF) (see section 2.2.1). A complete discussion and comparison of the effect of introducing this correction is included in [Olmedo et al., 2018]. Note that for the sake of simplicity the figures included here are the same than in [Olmedo et al., 2018]. So, the L3 and L4 products being distributed for the Mediterranean region by BEC are the ones labeled as *"corrected"*. However, in the rest of text and tables they are denoted as L3 and L4 products.

A.1 Comparison with Argo floats

Since we have used Argo floats for the generation of the time-dependent bias correction, this *in-situ* data set is not an independent source of SSS data to be used in our assessment. However, we include a comparison with Argo because our methodology uses the average in the full domain (North Atlantic and Mediterranean Sea) for removing the biases. A completely independent validation is presented in the next sections using moorings and on board thermosalinometer observations. We also want to highlight that the number of Argo surface measurements available for validation in the Mediterranean Sea is low (less than 50) before 2014. Thus, we analyze the SMOS SSS errors with respect to Argo in two different regions:

- Full domain (DOM), which includes part of the North Atlantic and the Mediterranean Sea and is defined in $[26°, 50°]\text{N}\times[-50°, 38°]\text{E}$
- The Mediterranean Sea (MED) which is defined by the following rectangle $[30°, 44°]\text{N}\times[-6°, 30°]\text{E}$.

Table 2 shows the statistics of the differences between SMOS and Argo SSS. The spatial distribution of the mean differences the different SMOS SSS products and Argo a major reduction of the bias in the Mediterranean basis (see Figures 5 and 6 in [Olmedo et al., 2018]).

A.2 Comparison with moorings data

In this section we have compared the time series of the 6 years of SMOS SSS with the time series provided by moorings located in the coastal Western Mediterranean , see their location in Figure 6).

The spatial and temporal scales corresponding to the mooring and satellite measured SSS are very different. While the satellite observations correspond to SSS over an area of $0.25° \times 0.25°$ (or $0.05° \times 0.05°$ in case of L4 products) and computed from an integrated time period of 9 days, the mooring measurements represent SSS values of a pinpoint location at hourly time steps. Therefore, in order to make more comparable both SSS data sets, we have applied a slide averaging window of 9 days to the mooring data. This smoothing technique does not solve completely the temporal scale differences in the comparison between mooring and satellite SSS (as we will discuss below).

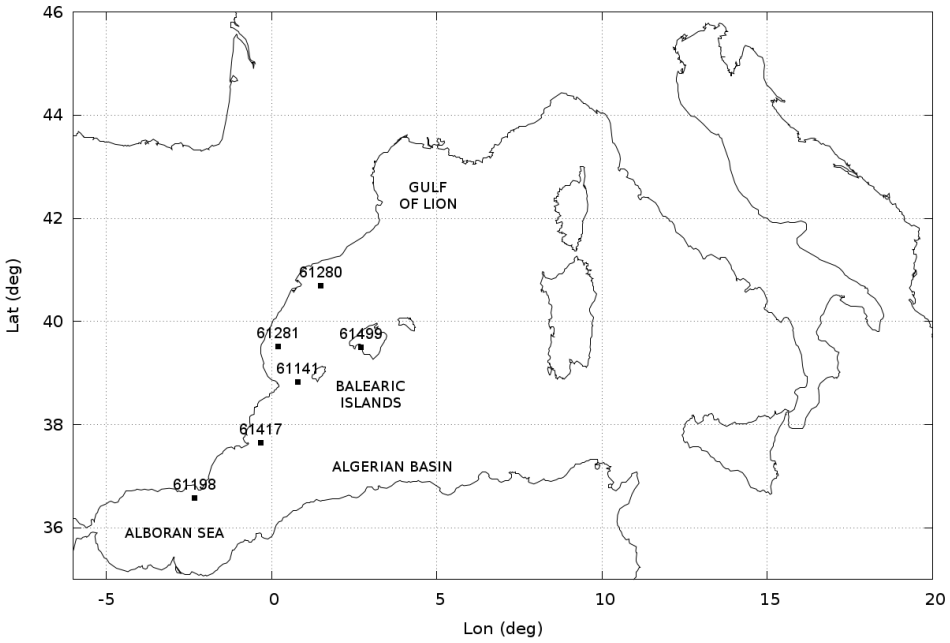


Figure 6: Location of the moorings that have been used for validation.

Table 3: Differences between SMOS and in situ SSS. For each one of the four SMOS products, the mean and the standard deviation are shown. In the first column the mooring ID is shown.

Mooring ID	L3		L4	
	mean	std	mean	std
61198	-0.20	0.31	-0.25	0.32
61417	-0.19	0.44	-0.26	0.42
61141	0.20	0.28	0.14	0.20
61499	0.20	0.27	0.23	0.15
61281	-0.22	0.38	-0.27	0.34
61280	-0.21	0.25	-0.22	0.22

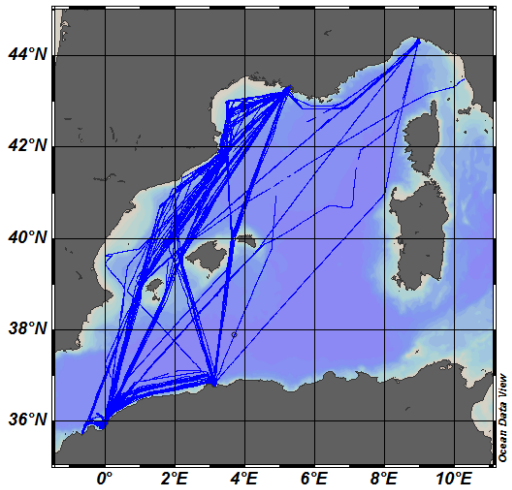
Figure 7 shows the time series corresponding to the four moorings anchored on the Spanish continental coast (moorings number 61198, 61417, 61281 and 61280, see their locations in Figure 6), where the *in-situ* salinity range is higher (> 1 PSU). In this case, although a slide averaging window of nine days has been applied to the mooring data, the time series of these moorings (blue lines) still present jumps with a fast dynamics that will not be captured by the satellite. When these jumps are persistent (for example January 2013 mooring number 6147), the satellite captures the jump. On the contrary, if the jump does not lead to a persistent jump (for example January 2012 mooring number 61198), the satellite will not capture the dynamics. In spite of the temporal scale limitations, both corrected products (grey and black points) are closer to the *in-situ* (blue points) than the non-corrected ones (yellow and green).

Table 3 shows the mean and the standard deviation of the differences between SMOS SSS and *in-situ* SSS. The moorings anchored in the Baleares (61141 and 61499) have a different behavior with respect to the rest of moorings. On one hand the residual bias is positive. On the other hand, the bias in the Palma bahia (61499) corresponding to the L4 corrected increases with respect to the bias of the L4 (as above mentioned).

A.3 Comparison with TRANSMED SSS

In-situ data from the TRANSMED thermosalinometer system, (http://www.mio.univ-amu.fr/?TRANSMED) on board the MV Marfret Niolon ([Taupier-Letage et al., 2014]) have been also used as an independent source of SSS data for validating the resulting SMOS products. These data cover from February 2012 to July 2014 and they are composed of SSS and SST acquired underway along repeated transects at ∼3m deep along the Western Mediterranean (see Figure 8).

Table 4 shows the statistics corresponding to the four different SMOS SSS products and the *in-situ* SSS. The time correction proposed in this work reduces the bias in both products L3 and L4 from -0.54 PSU to less than -0.1 PSU. On the other hand, a reduction of 0.12 (0.03) PSU in the standard deviation of the L3 (and L4 respectively) is also obtained.



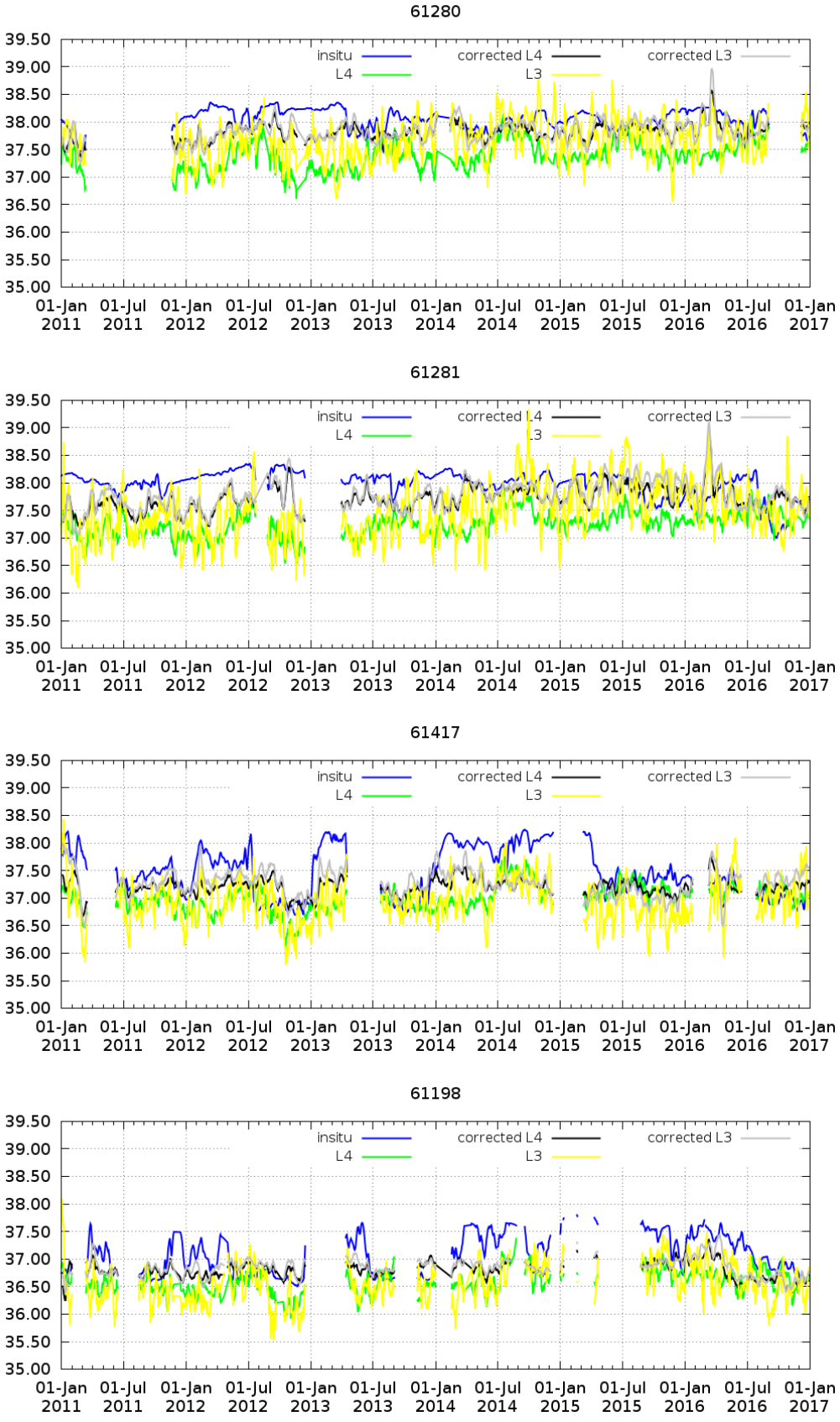


Figure 7: Moorings in the spanish continental Mediterranean coast. Time evolution of the salinity: in blue the in situ SSS; L3 SMOS SSS in yellow; L4 SMOS SSS in green; corrected L3 SMOS SSS in grey; and corrected L4 SMOS SSS in black. The location of the buoys are shown in Figure 6.

Table 4: Differences between SMOS and TRANSMED SSS values. For each one of the four SMOS SSS products the mean, the standard deviation and the total number of colocations between in situ and satellite used in the statistics are provided.

	L3	L4
Mean	-0.03	-0.09
Std	0.27	0.25
N_{meas}	97530	91113

In order to take a global look at the spatial biases of the SMOS SSS product with respect to TRANSMED SSS data, Figure 9 shows a latitudinal Hovmöller diagram of the different SSS data ordered in time. The time correction proposed here has strongly mitigated some particular events that contaminated the initial SMOS product. The Hovmöller plots show a similar behavior than the one observed by the TRANSMED SSS. In spite of the different spatial scales, which produce a more pixelated plot for SMOS, and in spite of a slight global fresh bias,also shown in the statistics of Table 4, the regional SMOS SSS products allow reducing the spatial noise and they also improve the depiction of the North-South SSS gradients. For instance, the fresher (blue) area moves further north around October 2013 in the TRANSMED SSS and both, L3 and L4 regional, SMOS products, whereas in March-April 2014, the salty water moves south.

In the northern part of the basin the comparison between SMOS SSS products and TRANSMED SSS is impaired mostly by the Rhone river plume. Its scales of variability in both space (along the shoreline or offshoreward) and time (hours under wind influence or heavy precipitating events) cannot be captured by SMOS. Further analysis on larger scales (northwestern part of the basin, season) would probably help assessing the freshwater discharges observed by SMOS in the Gulf of Lion.

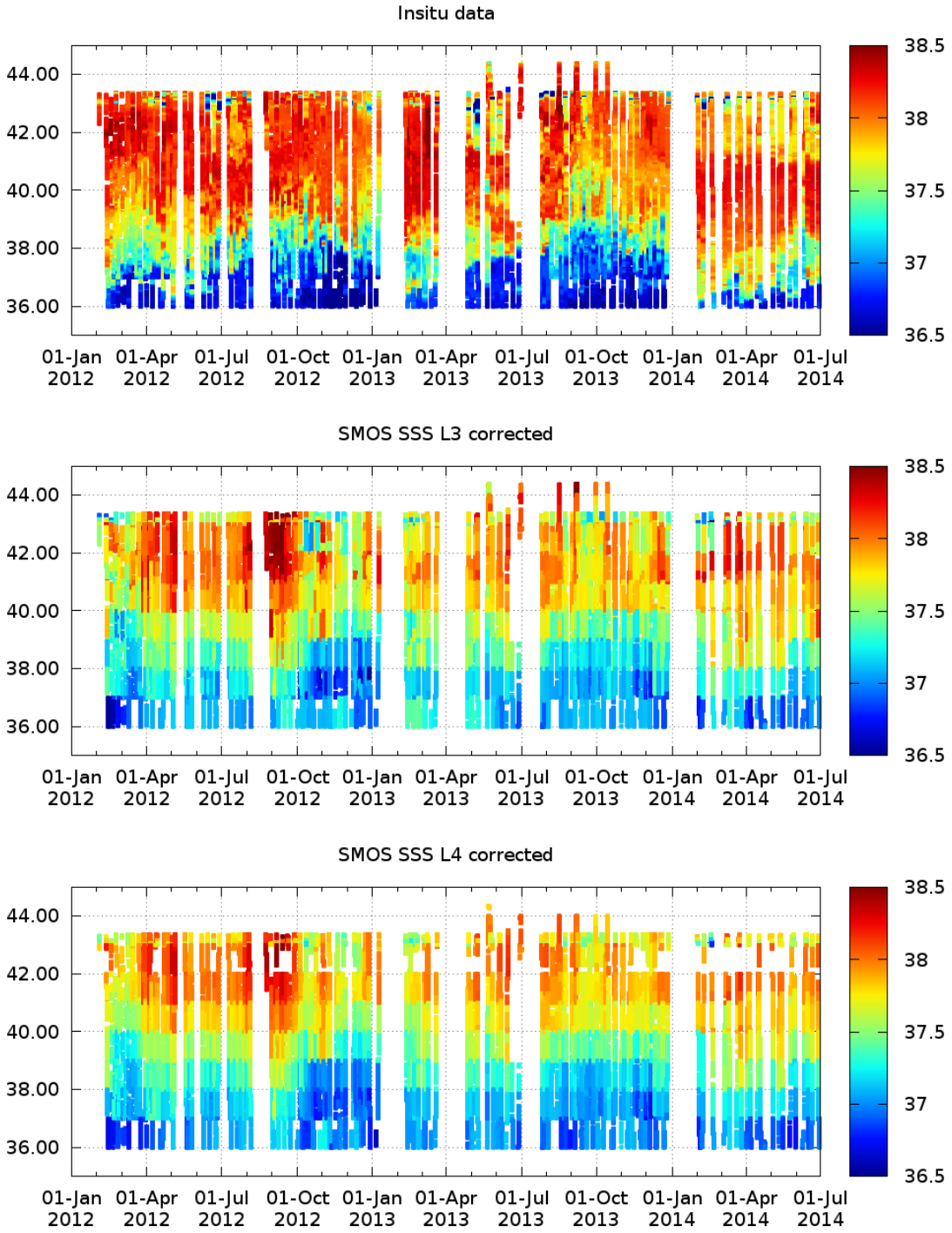


Figure 9: Latitudinal Hovmöller with the SSS values ordered in time corresponding to (from top to bottom) : TRANSMED SSS; the L3 corrected; and the L4 corrected SMOS SSS values, respectively

References

[Alvera-Azcárate et al., 2005] Alvera-Azcárate, A., Barth, A., Rixen, M., and Beckers, J.-M. (2005). Reconstruction of incomplete oceanographic data sets using Empirical Orthogonal Functions. Application to the Adriatic Sea surface temperature. *Ocean Modelling.*, 9:325–346. doi:10.1016/j.ocemod.2004.08.001.

[Alvera-Azcárate et al., 2009] Alvera-Azcárate, A., Barth, A., Sirjacobs, D., and Beckers, J.-M. (2009). Enhancing temporal correlations in EOF expansions for the reconstruction of missing data using DINEOF. *Ocean Science*, 5:475–485.

[Beckers and Rixen, 2003] Beckers, J.-M. and Rixen, M. (2003). EOF calculations and data filling from incomplete oceanographic data sets. *Journal of Atmospheric and Oceanic Technology*, 20(12):1839–1856.

[Font et al., 2010] Font, J., Camps, A., Borges, A., Martin-Neira, M., Boutin, J., Reul, N., Kerr, Y., Hahne, A., and Mechlenburg, S. (2010). SMOS: the challenging sea surface salinity measurement from space. *Proceedings of the IEEE*, 98:649.

[González-Gambau et al., 2017] González-Gambau, V., Olmedo, E., Martínez, J., Turiel, A., and Duran, I. (2017). Improvements on Calibration and Image Reconstruction of SMOS for Salinity Retrievals in Coastal Regions. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(7):3064 – 3078.

[González-Gambau et al., 2016a] González-Gambau, V., Olmedo, E., Turiel, A., Martínez, J., Ballabrera-Poy, J., Portabella, M., and Piles, M. (2016a). Enhancing SMOS brightness temperatures over the ocean using the nodal sampling image reconstruction technique. *Remote Sensing of Environment*, 180:202–220.

[González-Gambau et al., 2016b] González-Gambau, V., Turiel, A., Olmedo, E., Martínez, J., Corbella, I., and Camps, A. (2016b). Nodal Sampling: A New Image Reconstruction Algorithm for SMOS. *IEEE Transactions on Geoscience and Remote Sensing*, 54(4):2314–2328.

[Hernández-Carrasco et al., 2015] Hernández-Carrasco, I., Sudre, J., Garcon, V., Yahia, H., Garbe, C., Paulmier, A., Dewitte, B., Illig, S., Dadou, I., González-Dávila, M., and Santana-Casiano, J. M. (2015). Reconstruction of super-resolution ocean pco2 and airsea fluxes of co2 from satellite imagery in the southeastern atlantic. *Biogeosciences*, 12(17):52295245.

[Isern-Fontanet et al., 2016] Isern-Fontanet, J., Olmedo, E., Turiel, A., Ballabrera-Poy, J., and García-Ladona, E. (2016). Retrieval of Eddy dynamics from SMOS sea surface salinity measurements in the Algerian Basin (Mediterranean Sea). *Journal of Geophysical Research Letters*, 43:6427–6434.

[Isern-Fontanet et al., 2007] Isern-Fontanet, J., Turiel, A., García-Ladona, E., and Font, J. (2007). Microcanonical multifractal formalism: Application to the estimation of ocean surface velocities. *Journal of Geophysical Research: Oceans*, 112(C5):2156–2202.

[Kerr et al., 2010] Kerr, Y., Waldteufel, P., Wigneron, J.-P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M.-J., Font, J., Reul, N., Gruhier, C., Juglea, S., Drinkwater, M., Hahne, A., Martin-Neira, M., and Mecklenburg, S. (2010). The SMOS mission: new tool for monitoring key elements of the global water cycle. *Proceedings of the IEEE*, 98(5):666–687.

[McMullan et al., 2008] McMullan, K. D., Brown, M., Martin-Neira, M., Rits, W., Ekholm, S., Marti, J., and Lemanczyk, J. (2008). SMOS: The Payload. *Geoscience and Remote Sensing, IEEE Transactions on*, 46(3):594–605.

[Nieves et al., 2007] Nieves, V., Llebot, C., Turiel, A., Solé, J., Garca-Ladona, E., Estrada, M., and Blasco, D. (2007). Common turbulent signature in sea surface temperature and chlorophyll maps. *Geophysical Research Letters*.

[Olmedo et al., 2016] Olmedo, E., Martínez, J., Umbert, M., Hoareau, N., Portabella, M., Ballabrera-Poy, J., and Turiel, A. (2016). Improving time and space resolution of smos salinity maps using multifractal fusion. *Remote Sensing of Enviroment*.

[Olmedo et al., 2017] Olmedo, E., Martnez, J., Turiel, A., Ballabrera-Poy, J., and Portabella, M. (2017). Debiased non-bayesian retrieval: A novel approach to smos sea surface salinity. *Remote Sensing of Environment*, 193:103 – 126.

[Olmedo et al., 2018] Olmedo, E., Taupier-Letage, I., Turiel, A., and Alvera-Azcrate, A. (2018). Improving smos sea surface salinity in the western mediterranean sea through multivariate and multifractal analysis. *Remote Sensing*, 10(3).

[Taupier-Letage et al., 2014] Taupier-Letage, I., Bachelier, C., and Rougier, G. (2014). Thermosalinometer TRANSMED, Marfret Niolon, definitive data set. *SEDOO OMP*. http://dx.doi.org/10.6096/MISTRALS-HyMeX.1127.

[Turiel et al., 2008] Turiel, A., Solé, J., Nieves, V., Ballabrera-Poy, B., and García-Ladona, E. (2008). Tracking oceanic currents by singularity analysis of Microwave Sea Surface Temperature images. *Remote Sensing of Environment*, 112(5):2246 – 2260. Earth Observations for Terrestrial Biodiversity and Ecosystems Special Issue.

[Umbert et al., 2014] Umbert, M., Hoareau, N. Turiel, A., and Ballabrera-Poy, J. (2014). New blending algorithm to synergize ocean variables: the case of SMOS sea surface salinity maps. *Remote Sensing of Environment 146, pp. 188-200*.

[Zweng et al., 2013] Zweng, M. M., Reagan, J. R., Antonov, J. I., Locarnini, R. A., Mishonov, A. V., Boyer, T. P., Garcia, H. E., Baranova, O. K., Johnson, D. R., Seidov, D., and Biddle, M. M. (2013). *World Ocean Atlas 2013, Volume 2: Salinity*. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 74, 39 pp.